

REMARKS

Claims 30-41 and 45-62 are currently pending in the application. Claims 42-44 are being canceled without prejudice or disclaimer. Claims 30, 32, 33, 34, 35, 36, 38, 40, 45, 51, 52, 53, 56, 57, and 58 are currently amended in this response and Claims 60-62 are newly added claims.

I. RESPONSE TO CLAIM OBJECTIONS

Claims 30, 34, 38, 42, 45, 49, 53, and 57 are objected to in that the Examiner stated the limitation “a user defined minimum number of words expressed in *a* first language” should be “a user defined minimum number of words expressed in *the* first language”. Applicant agrees with the Examiner, and apologizes for the antecedent basis error. The claims are herein amended to account for this objection. Moreover, claims 30, 34, 38, 45, 53, and 58 have been amended to account for the minor change related to the phrase “analysis,” which should be “analyzing” as set forth in those claims. Finally, claims 30, 32-34, 36-38, and 40 are amended to note that the target language word phrases can be manipulated in any order and may be ranked by ascending number of second-language words that are not direct translations of words in the original first language query, i.e. by the fewest number of inserted words that are not translations of the original query.

II. RESPONSE TO REJECTION UNDER 35 U.S.C. § 102

Claims 30, 32 – 34, 36 – 38, and 40 – 44 were rejected as being anticipated by Grefenstette under 35 U.S.C. § 102. Applicant respectfully disagrees with this rejection, as Grefenstette does not teach all elements of the asserted claims, as described herein.

Generally, Grefenstette teaches a form of example based translation of very short noun phrases wherein certain language-specific ordering rules are first applied manually, and then applies its translation to candidate two word noun phrases. Grefenstette ultimately identifies a translation from the translated phrase candidates based on the frequency of the candidates on the World Wide Web.

Unlike the invention disclosed by the referenced claims, Grefenstette does not teach a method for automatically translating between languages without the use of word ordering rules between languages and does not teach a method for translating arbitrarily long strings, in particular strings larger than two word noun phrases. Grefenstette requires that the translations occur in a specific order, which is totally different from the order-freedom contemplated by the invention disclosed by the referenced claims. The invention allows for returning of any word orderings – since a true translation across languages may invert or otherwise permute the word ordering. This is key to producing good translations and it is missing from Grefenstette.

Thus, before searching for phrasal translations among the candidates in the target language (in this case English), Grefenstette requires as a first step the ordering of potential translations into their correct, proper noun phrases based on grammatical ordering rules to form the ultimate translation. For example, when Grefenstette searches for the candidate translations for the noun phrase "groupe de travail" a manual determination is made to "reverse" the order of the words "groupe" and "travail" for translation of the phrase into English. That is, the manual ordering of Grefenstette looks at only the possible English translation for the word "groupe" in the second position, and the English translations for the word "travail" in the first position of the candidate English phrases.

As described by Grefenstette:

Combining the possible translations of groupe de travail into all twenty-one possible noun phrases **creatable by simply re-ordering the nouns and concatenating them to form English phrases and then submitting these phrases to this** Web browser yields, in Table 2, the actual occurrence statistics in the web pages indexed by this browser.

[Page 3, paragraph 7](emphasis added). An examination of Table 2 in Grefenstette shows that only fifteen possible combinations (not twenty-one as Grefenstette mistakenly says) of translations of "groupe" in the second position and "travail" in the first position were examined: three possible translations for

“travail” and five possible translations for “groupe.” Grefenstette makes this clear when it says:

Take the compositional French noun phrase *groupe de travail*. In the Oxford-Hachette French-English dictionary, the French word *groupe* can be translated by the English words cluster, group, grouping, concern and collective. The French word *travail* can be translated by the English words work, labor, or labour. The native translator has five (from *groupe*) times three (from *travail*) possible ways of translating *groupe de travail*.

[Page 2, paragraph 6] (Emphasis Added) As shown in Table 2, thirteen of the combinations existed at least once on the web (the other two did not occur as described by the note under Table 2). The order dictated by the re-ordering process yields only fifteen phrasal translation candidates.

Moreover, in the example of translating the noun phrase “*groupe de travail*”, many of the two word phrases in the *reverse* order exist on the web but would not be correct translations. That is, if Grefenstette were to allow translation in any order, there would actually be thirty possible candidates (the three occurrences of “*travail*”, times the five occurrences of “*groupe*”, times two for accounting for all potential ordering.

The fact that Grefenstette says that the five translations for *groupe* multiplied by the three translations for *travail* gives the universe of possibilities implies that they are only going to order the words one way (i.e., the translations for *groupe* in the second position and translations for *travail* in the first position). If they were to look at all possible word orders for the translations

there would have been five times three times two (to give both orders for each combination) possible phrasal translations.

The present invention disclosed in the referenced claims do not account for any word particular word order; indeed, Grefenstette teaches away from analyzing *all* word orders in a phrase. Because Grefenstette identifies the most frequent occurrences of candidate translations, eliminating those translations with ordering rules is crucial for the process in Grefenstette to be effective. For example, the two-word phrase "collective labor" (which is one combination with the translation for groupe in the first position and travail in the second position) has many appearances on the web but was not considered because the ordering rules in Grefenstette made it not valid as a candidate to be tested.

It should also be noted that Grefenstette requires pre-generating all combinations of translated words and then simply matches them against the target corpus. Whereas this may work for two-word noun phrases, it totally fails at translating longer phrases, clauses and whole sentences. For instance if a 12 word sentence were to be translated in Grefenstette's pre-generation of ordered strings, and each word had an average of 5 translations (which is typical), then there would be 5 exponentiated to the 12th power strings to generate, which is approximately $\frac{1}{4}$ *billion* sequences, and therefore computational not viable in reducing the invention to practice. In contrast, the invention does not require pre-generation of ordered strings, but instead searches the second-language

corpus for strings that contain the query words (2 or 5 or 12 or any number of query words). Containment does not imply order, and hence it is much more efficient to compute – no pre-generation of a quarter billion strings *per query*, to enforce ordering in required in the invention. Hence the invention is scalable to translating any text, whereas Grefenstette is capable only of operating on very short sequences, such as 2-word noun phrases. This is a crucial difference.

In addition, Grefenstette only applies to situations where there is an ordering of the words based on knowledge of the relationship between the source and target language to provide possible noun phrases in the target language are the tables reflecting results in translating German noun phrases into English and Spanish noun phrases into English. In every case, the order of the translations of the two words always appears in the same order. Grefenstette goes on to say that:

This experiment could also be made more subtle by generating more varied syntactic forms (such as A of B) or through a more intelligent use of morphological variants, without modifying the way that the available Web browser indexes its pages (emphasis added).

[Page 7, paragraph 16] (Emphasis added) Here Grefenstette points out that the invention can be expanded by generating other syntactic forms such as the passive voice for noun phrases. It is clear that Grefenstette needs to apply

ordering rules to form possible noun phrase translations in the target language for their “highest frequency on the web” approach to work.

In addition, Grefenstette does not anticipate the referenced claims because Grefenstette ordering rules would not effectively work as the phrases to be translated go beyond two word phrases. Although it's clear that Grefenstette needs ordering rules to work effectively even for two word phrases, the word order issue gets exponentially larger (i.e., more possible orders in the target) as the translation query is expanded to three or more words. The referenced claims of the present invention uses word string phrases of any size and allows the words to reorder naturally based on the appearance of the phrase on the web.

Finally, the referenced claims of the present invention allow for certain words in the target language translation candidate to be words that are not translations of any of the source language words. This is clear because the present invention looks for “a Target Language word string of user-defined maximum length” that contains a “user-defined minimum number” of source language word translations. The process in Grefenstette does not allow for “extra words” that are not translations of a source word in the query.

Applicant's invention uses word strings of any length and looks at all possible orderings of the words to generate translation candidates. If Grefenstette allowed for such a broad approach, their invention would not produce accurate results because the sole criteria for choosing a result using

Grefenstette is the frequency of occurrence of the candidate translation on the web. In contrast, the present invention intentionally over-generates candidate word string translations when searching for matches in a corpus of documents like the web.

In light of the above arguments, applicant has amended the claims to make this point clear by explicitly stating that the words in the translation candidate can appear in any order. Support for this amendment is shown in the specification of the present invention: "A Target Language word string of user-defined maximum length will qualify if it contains any combination, found in any order, of the user-defined minimum number of candidates generated by the different Source Language words", see paragraph [0226]

III. RESPONSE TO REJECTION UNDER 35 U.S.C. § 103

The examiner has also rejected claims 31, 35, and 39 under 35 U.S.C. § 103(a) as being patentable over Grefenstette in view of U.S. Patent No. 5,311,429 ("Tominaga"). Applicant respectfully disagrees, and notes that the arguments in the section above, detailing the differences between Grefenstette and the claimed invention, would not make claims 31, 35, and 39 obvious to one of ordinary skill in the art.

In addition, Tominaga would not, in combination with Grefenstette, make the invention of claims 31, 35, and 39 obvious because Tominaga uses sentence

generation rules based on the co-occurrence of a word in two different sentences. That is, Tominaga requires rules that would eliminate words or word phrases from the overlap process to obtain a consistent resultant phrase. For example, the examiner discusses the two phrases "the girl is sitting in the front row" and "the girl is my niece", which in turn generate a third sentence "the girl sitting in the front row is my niece". It's clear that rules are needed to generate that sentence since the word "is" is eliminated from the part of the third sentence that is generated from sentence one: "the girl IS sitting in the front row". In the invention of claims 31, 35, and 39, overlapping portions are combined in a way that matches the phrases to yield a third, candidate translation without application of any ordering rules, such as: "the girl in the front" and "in the front row" yielding "the girl in the front row." The overlapping word-string "in the front" are overlapped to return the candidate phrase, which is different that the process described in Tominaga.

IV. DEPENDENT CLAIM OBJECTIONS


The examiner states that claims 45-56 are objected to as being dependent upon a rejected base claim, but that the dependent claims would be allowed if rewritten in independent form to include all the limitations of the base claim and any intervening claims. Applicant appreciates the Examiner's review of these claims, but believes that the amended base claims make those base claims

allowable over the prior art, and hence moot this objection. To the extent the examiner disagrees with applicant's arguments regarding Grefenstette and Tominaga, applicant will rewrite these objected-to claims as specified.

If necessary to effect a timely response, this paper should be considered as a petition for an Extension of Time sufficient to effect a timely response, and please charge any deficiency in fees or credit any overpayments to Deposit Account No. 05-1323 (Docket #101205.55175C6).

Respectfully submitted,

September 29, 2008



Michael J. Songer
Registration No. 39,841

CROWELL & MORING LLP
Intellectual Property Group
P.O. Box 14300
Washington, DC 20044-4300
Telephone No.: (202) 624-2500
Facsimile No.: (202) 628-8844
MJS:alw